УДК 004.032.26 (06)

ИНФОРМАЦИОННАЯ НЕЙРОСЕТЕВАЯ СИСТЕМА ДЛЯ ИНТЕЛЛЕКТУАЛЬНОГО АНАЛИЗА ДАННЫХ

Н.В. Замятин, В.П. Максимов, Н.В. Платонов, М.Н. Тарасевич

Томский университет систем управления и радиоэлектроники
E-mail: zam@fet.tusur.ru

Изложена методика нейросетевого анализа данных (data mining). Показано, что применение нейронных сетей Кохонена позволяет эффективно выделять группы связанны данных и определять между ними закономерности. Разработана информационная система для геофизической предметной области.

Развитие технологии интеллектуального анализа данных (ИАД) вызвано рядом объективных факторов. Главные из них: накопление большого количества данных в различных предметных областях и скорость накопления информации намного превысила скорость ее обработки. Только научные учреждения за один день записывают информации примерно на 1 терабайт (по данным аналитического отдела американской компании GTE). При этом наука является не самым большим источником данных, и существуют огромные базы данных в сфере коммерции, энергетики, геологии, медицине, управляющих структурах оргсистем.

На сегодняшний день технология ИАД содержит множество различных подходов к обнаружению знаний. Каждый из них имеет свои преимущества и недостатки. При этом выбор конкретного подхода определяется спецификой предметной области и организацией данных [1]. Целесообразно применение ИАД для выявления знаний в большом количестве данных сейсмической разведки полезных ископаемых.

При обработке результатов в сейсморазведке для построения прогностических моделей традиционно используется аппарат многомерной линейной регрессии. Его использование имеет следующие сложности и ограничения:

- Ограничение классом линейных зависимостей. Искомая прогностическая модель сразу предполагается линейной. Хотя и возможно использование нелинейного преобразования независимого параметра, эта процедура носит достаточно произвольный характер, и итоговая модель все равно будет линейно зависеть от преобразованных параметров.
- Сложность выделения влияющих параметров. Из-за большого количества динамических параметров некоторые из них даже не рассматриваются как кандидаты на участие в модели. С другой стороны, в модель, согласно требованиям классической статистики, включаются только независимые параметры. Однако, зависимые параметры, в совокупности, могут также нести ценную информацию о целевом параметре.
- Зашумленность входной информации. Сейсморазведка района проводится в течение нескольких лет различными исследовательскими партиями, с использованием различного оборудования и т. п. Поэтому, зачастую в данных разведки встречаются выбросы – резко нетипичные значения. Эти выбросы значительно влияют на строимые линейные модели.

Аппарат нейронных сетей свободен от перечисленных недостатков: нейросети могут аппроксимировать любую непрерывную функцию, автоматически проводят анализ чувствительности влияния входных параметров на результат, устойчивы к шуму в исходных данных.

В качестве основных функциональных требований к интеллектуальным нейросетевым системам можно выделить:

- совместимость форматов хранения информации с наиболее распространенными средствами табличной обработки данных (MS Access, MS Excel).
- возможность обработки не только числовой, но и текстовой информации,
- возможность нормализации исходных обучающих данных различными способами,
- классификация обучающих данных посредством обучения нейросети.
- визуализация полученных карт различными способами.

Выделенные функциональные требования упорядочены в соответствии с жизненным циклом процесса ИАД, что позволяет создать информационную систему интеллектуального анализа данных (ИСИАД), предназначенную для решения задач кластеризации и классификации разнородной информации. Способом классификации данных, реализованным в системе, является нейронная сеть Кохонена.

Нейронная сеть Кохонена (самоорганизующаяся карта Кохонена) решает задачи классификации многомерных векторов. Достоинством сети, по сравнению с другими алгоритмами, является легкость визуализации и интерпретации полученных результатов. Обучение сети проходит без учителя, только на основе выборки входных данных (так называемое неуправляемое обучение).

Различные типы визуализации обученной сети позволяют легко выявить структуру входной информации: унифицированная матрица расстояний

отображает кластерную структуру данных, график компонентов позволяет установить форму зависимости входных параметров, плоскость компонентов и карта попаданий отражают распределение входных параметров.

Структурная схема ИСИАД приведена на рис. 1.

Модуль предобработки выполняет создание и хранение таблицы обучающих данных. Дополнительно модуль обеспечивает нормализацию таблицы обучающих данных. В качестве средства хранения была выбрана СУБД MS Access, являющаяся составной частью модуля.

Функция «Импорт данных» в ИСИАД реализуется средствами СУБД. Поэтому важным критерием выбора СУБД становится развитость ее средств обмена данными.

Функция «Создание классификаторов» сопоставляет текстовым значениям категориальных (перечислимых) данных числовой код. После такого сопоставления все атрибуты принимают только числовые значения, и данные являются подготовленными для использования в алгоритме обучения.

Нормализация данных осуществляется по указанию пользователя, масштабированием значений каждого параметра в диапазон [0;1]. Нормализуются уже подготовленные данные.

Модуль обучения нейросети реализует итеративный алгоритм обучения карты Кохонена.

Функция «Настройка параметров» данного модуля позволяет настроить конфигурацию самой карты (размеры карты и тип), а также параметры ее обучения (количество итераций, способ инициализации и др.).

Функция «Обучение» непосредственно позволяет обучить сеть Кохонена. Для доступности данной функции предварительно должны быть подготовлены исходные данные.

Расчет качества классификации может быть выполнен только после обучения нейросети.

После обучения карты Кохонена дополнительно может быть выполнена ее кластеризация алгоритмом *k*-средних. Алгоритм *k*-средних может применяться непосредственно к исходным данным. Однако, данный алгоритм имеет ряд недостатков и вычислительно сложен. К тому же, для классификации новых данных требуется новая прогонка алгоритма.

Модуль визуализации является одним из важнейших модулей, так как анализ полученной карты строится на различных способах ее графического отображения.

Данный модуль реализует четыре способа отображения карты Кохонена: унифицированная матрица расстояний, карта попаданий, плоскость значений некоторого компонента, график компонентов.

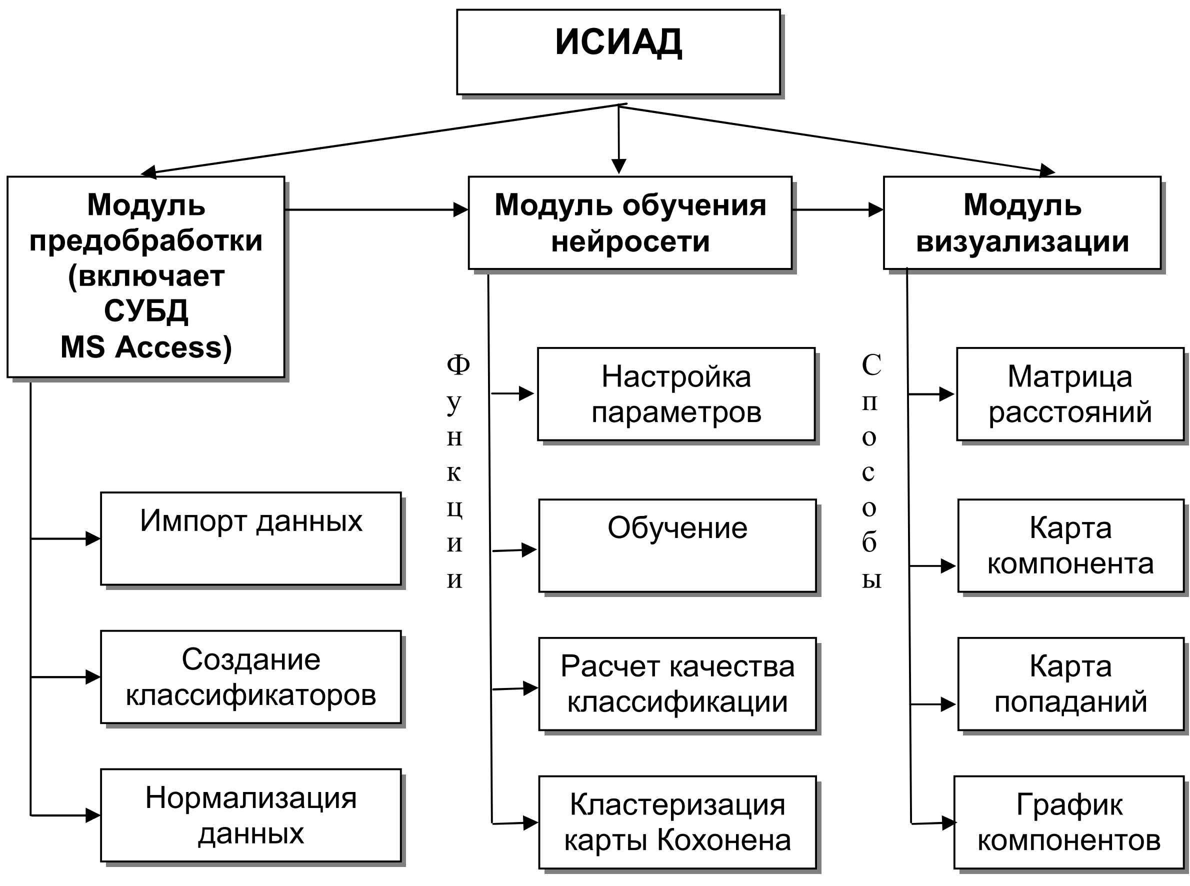


Рис. 1. Структурная схема ИСИАД

Визуализация карты требует дополнительных промежуточных вычислений. Результаты этих вычислений сохраняются в базе данных для последующего использования.

В качестве исходных данных для тестирования информационной системы использованы данные сейсмической разведки. Процесс сейсмической разведки заключается в проведении последовательных взрывов зарядов на местности через определенные расстояния (50, 100 и т. д. метров). После каждого взрыва, установленные датчики фиксируют параметры взрывной волны, отраженной от геологических пластов. Линия, вдоль которой проводятся взрывы, называется профилем.

Для построения прогностических моделей карты параметров дискретизируют и используют сетки параметров – значения параметра в узлах регулярной сетки (шаг сетки – 100 м).

Основной задачей сейсморазведки является прогнозирование значений параметров, влияющих на содержание нефти в пластах породы, вдоль карты района. Такими параметрами являются: пористость и проницаемость горизонта. Исходными данными для прогноза являются сетки структурных и динамических параметров, соотнесенные с пробами пластов.

Обработка полученной информации позволяет выделить различные акустические характеристики пластов (горизонтов). Причем количество этих параметров может быть достигать до 200 и более. Примеры параметров: время прихода отраженной от горизонта волны, средняя энергия, амплитуда волны, фаза и т. д.

Параметры делятся на структурные, относящиеся к профилю разведываемой поверхности, (время прихода волны, глубина) и динамические, описывающие отраженную взрывную волну (амплитуда, энергия).

Значения параметров измеряются вдоль взрывных профилей. Затем проводится интерполяция параметров между профилями. В итоге получается карта параметра – графическое изображение распределения значения параметра по равномерной сетке. Всего, в виде сеток, получены значения 26 параметров. Все они участвовали в обучении нейросети. Задачей являлось установление зависимостей между параметрами и степени их влияния на целевой параметр пористости.

Пример карты параметра приведен на рис. 2.

Пунктирными линиями на рисунке показаны профили взрывов. Значения входных параметров могут быть определены не в каждой точке сетки, например, из-за отсутствия вблизи данной точки

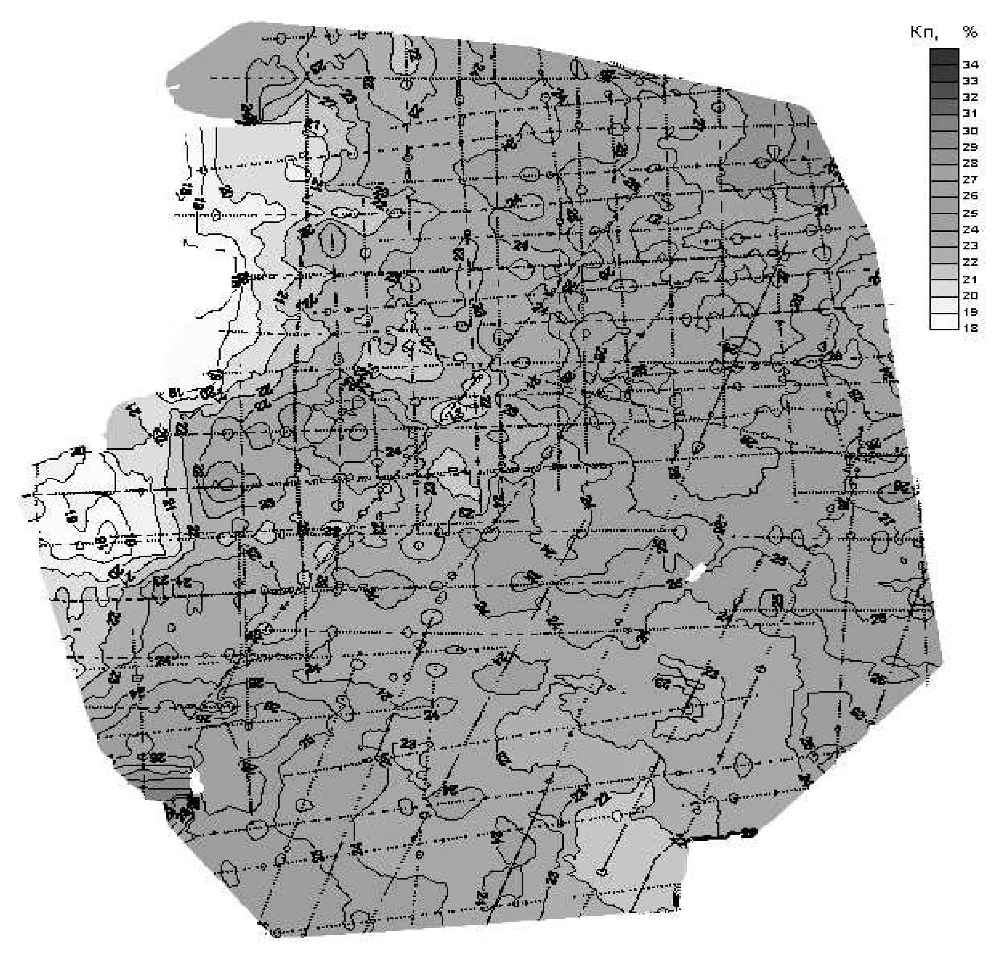


Рис. 2. *Карта параметра «Пористость горизонта b9», b9 – значения сетки*

взрывного профиля. Поэтому в исходной таблице данных содержатся пропущенные значения. Строки, содержащие пропущенные значения, были удалены из таблицы (очистка данных). После очистки осталось 54509 записей.

Перед обучением нейросети процесс подготовки данных заключается в присвоении числовых значений текстовым параметрам и (если задан соответствующий параметр нейросети) нормализации полученной таблицы.

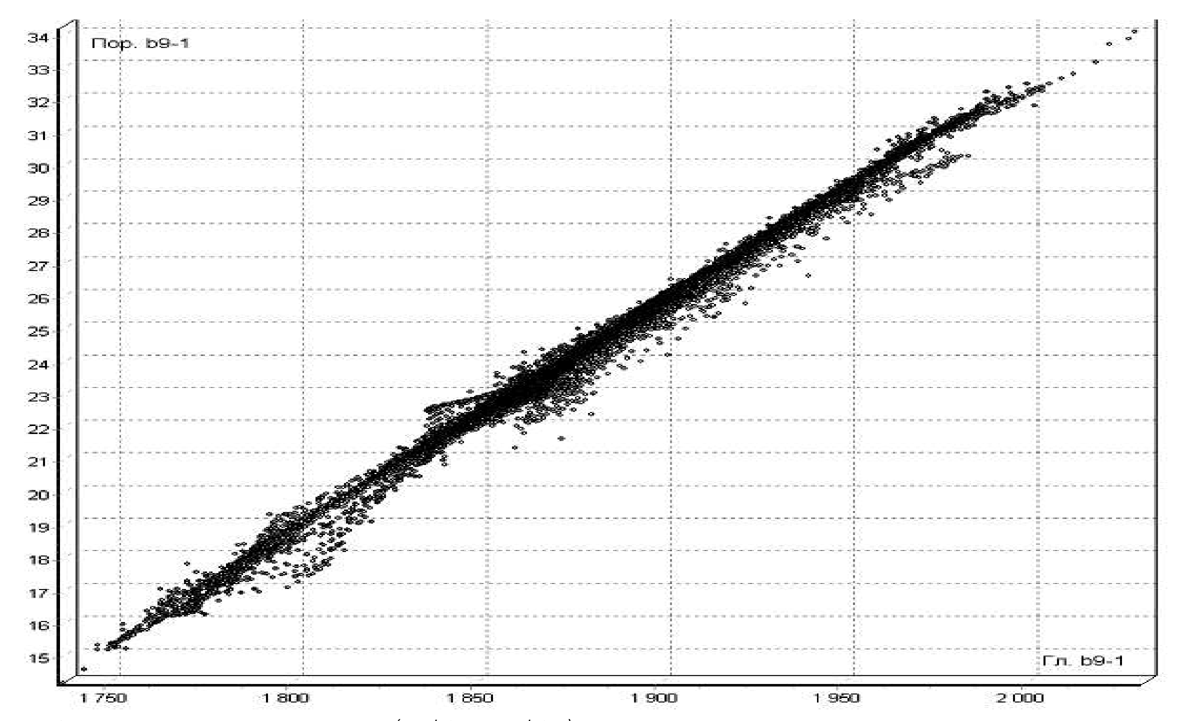


Рис 3. *График компонентов для пары (Гл. b9-1, Пор b9-1)*

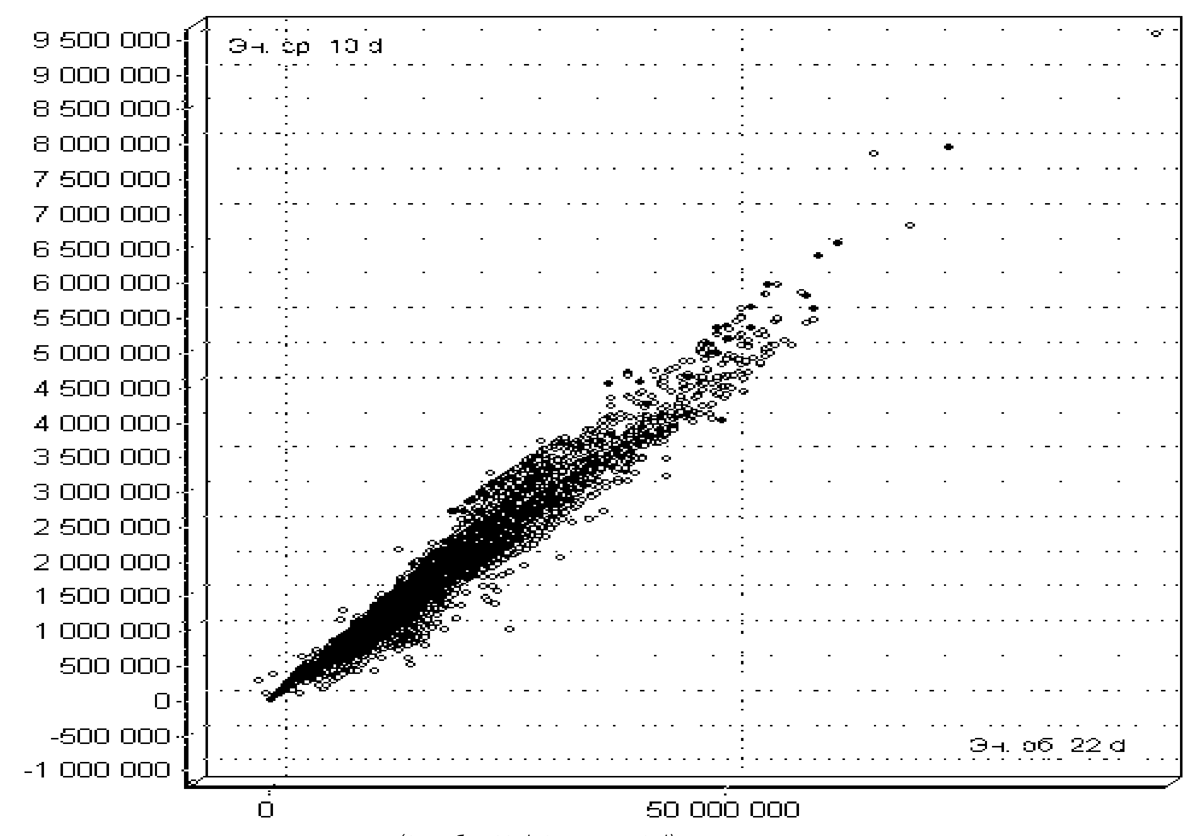


Рис 4. *График компонентов для пары (Эн. общ. 22 d, Эн. сред. 10 d)*

Для поиска зависимостей между параметрами используют визуализированные плоскости компонентов. Для этого в ИСИАД формируются плоскости всех компонентов и осуществляется их перегруппировка, располагающая сходные плоскости рядом. Такое расположение позволяет легко выделить зависимые компоненты входных векторов, поскольку для них в одинаковых местах карты будут находиться схожие шаблоны карты. Выявлено четыре группы коррелирующих компонентов. Установление конкретных форм зависимостей проводится при помощи графиков компонентов.

На рис. 3 приведен график компонентов для пары данных (Глубина b9-1, Пористость b9-1).

Анализ рисунка позволяет выявить четкую линейную зависимость между этими параметрами. Установление параметров данной зависимости может быть проведено при помощи любого из статистических методов. График компонентов для пары (Энергия. общая. 22 d, Энергия. средняя. 10 d) приведен на рис. 4.

Между данными параметрами также существует линейная зависимость. График компонентов для пары (Общая. абсолютная. амплитуда. b9-1, Энергия. общая. b9-1) приведен на рис. 5.

Видно, между параметрами энергии и амплитуды существует зависимость, близкая к квадратичной.

Анализ графиков других пар компонентов позволяет установить факты существования между ними линейных зависимостей. Данная группа параметров является сильно коррелирующей между собой. Поэтому, при построении модели численного прогноза в число значимых параметров следует включать только один из них.

Четвертая группа коррелирующих параметров интересна тем, что о данной зависимости не было ничего известно до начала процесса ИАД. На рис. 6 приведен график компонентов для пары (Дискрета 44 IIa, Проницаемость b9-12 16ob).

Зависимость между компонентами подобна линейной, однако разброс точек от основной линии достаточно велик. Это говорит о том, что на параметр проницаемости влияют также и другие параметры, не вошедшие в анализ.

Проведенный анализ не выявил во входной информации кластерной структуры. Это связано с тем, что все параметры являются численными характеристиками и имеют непрерывный числовой диапазон изменения, причем большинство из них независимо. В связи с этим, нельзя было сравнить распределение значений компонентов по различным кластерам, сравнить «населенность» кластеров (используя карту попаданий) и на этой основе дать описание типичных представителей кластера.

В то же время проведенный интеллектуальный анализ данных позволил:

- получить наглядное представление о структуре входной информации.
- выделить группы зависимых компонент и выявить тип зависимостей между компонентами.

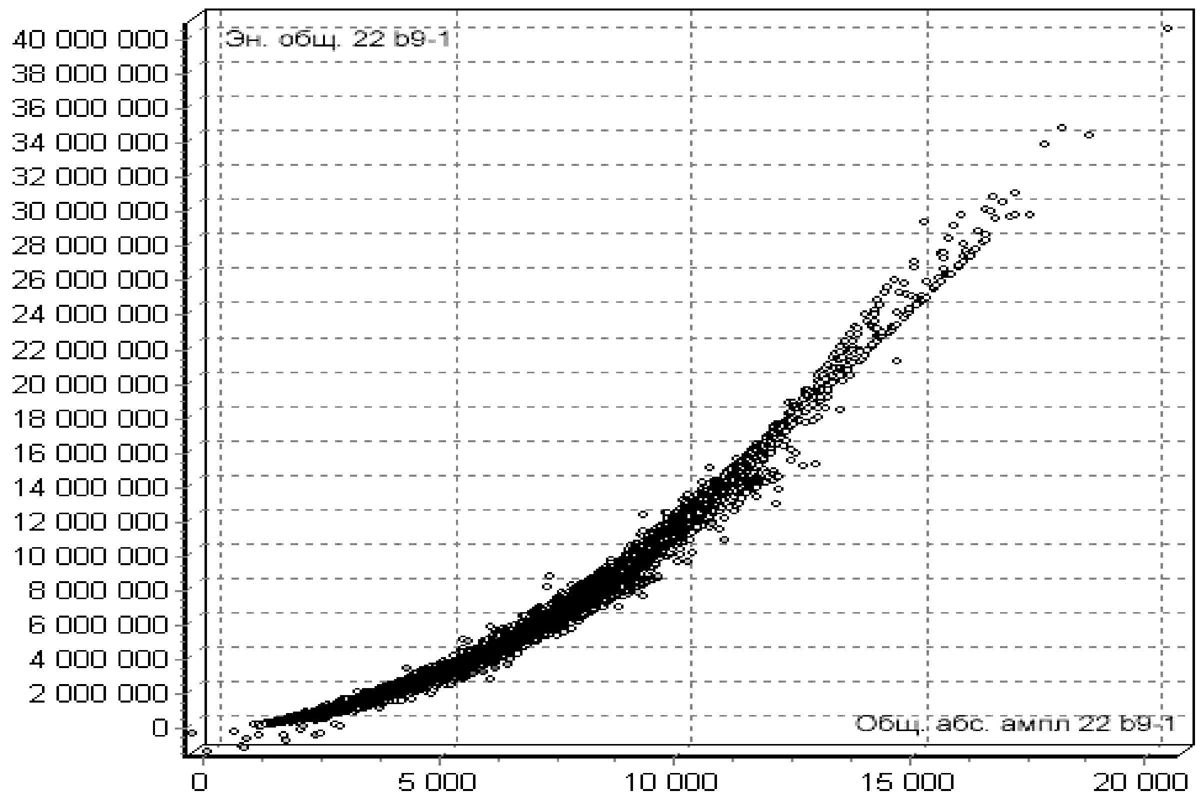


Рис 5. *График компонентов для пары (Общ. абс. ампл. 22 b9-1, Эн. общ. 22 b9-1)*

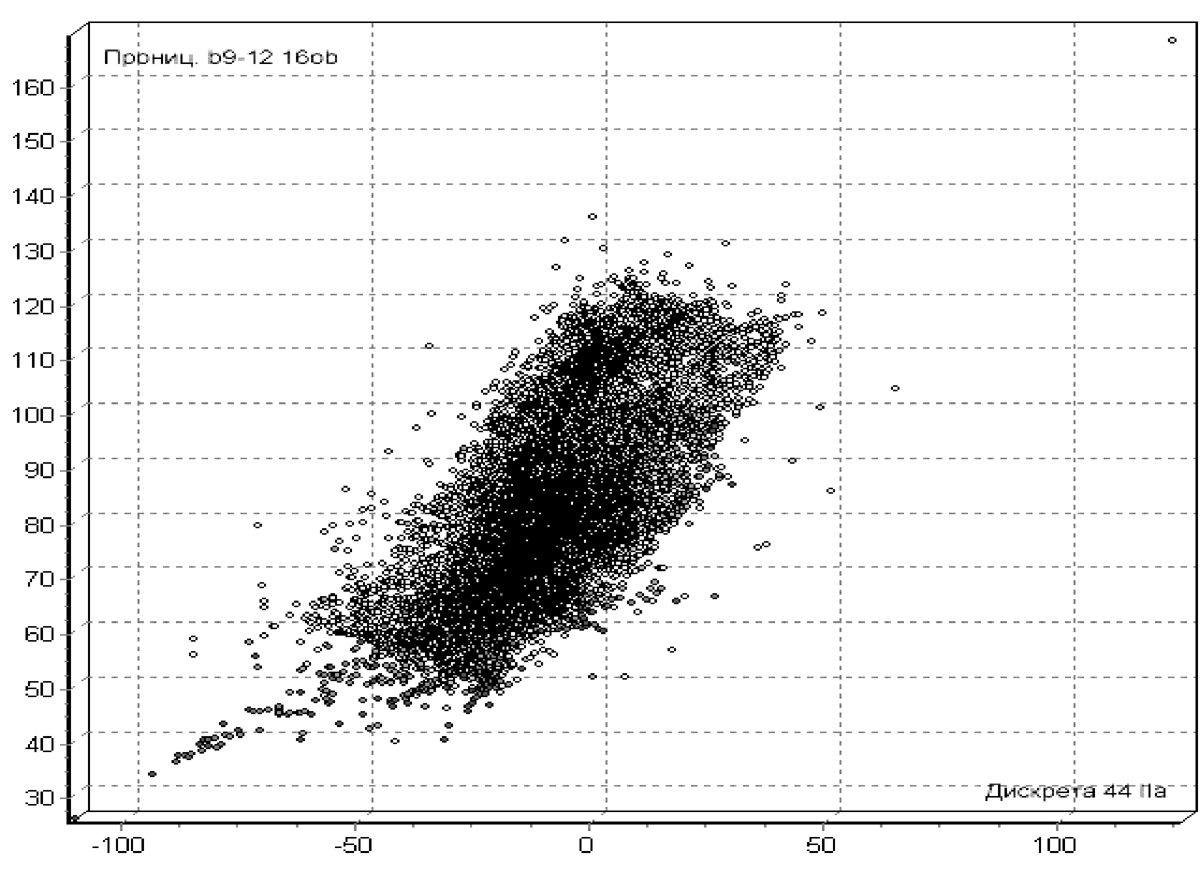


Рис 6. *График компонентов для пары (Дискрета 44 IIa, Прониц b9-12 16ob)*

- сделать вывод об отсутствии кластерной структуры входных данных, что обусловлено независимостью большей части параметров.
- выделить четыре группы коррелирующих параметров, содержащие от двух до восьми компонент.
- установить формы зависимостей параметров внутри групп.

Полученная информация может быть использована при построении числовых моделей прогнозирования целевых параметров, либо с помощью классических статистических методов, либо с помощью искусственных нейронных сетей с обратным распространением ошибки, что открывает большие возможности по интерпретации полученных результатов. Определение группы параметров, коррелирующих с целевой функцией, позволяет сразу выделить значимые компоненты, которые должны войти в числовую модель прогноза, и, тем самым, сократить размерность задачи.

Разработанная информационная система не привязана к конкретной предметной области. Это позволяет аналитику применять ее в любой области, после соответствующей предобработки накопленной в ней информации.

СПИСОК ЛИТЕРАТУРЫ

1. Корнеев В.В., Гареев А.Ф., Васютин С.В. Базы данных. Интеллектуальная обработка информации. – М.: Нолидж, 2000. – 352 с.
2. Уоссермен Ф. Нейрокомпьютерная техника: теория и практика. Пер. с англ. – М.: Мир, 1992. – 127 с.
3. Александров В.В. Интеллект и компьютер. – СПб.: Анатолия, 2004. – 285 с.
4. Львов В.В. Создание систем поддержки принятия решений на основе хранилищ данных // Системы управления базами данных. – 1997. – № 3. – С. 30–40.

УДК 004.93'12

МЕТОДИКА ЛОКАЛИЗАЦИИ ИЗОБРАЖЕНИЯ ЛИЦА ДЛЯ СИСТЕМ ВИДЕОКОНТРОЛЯ НА ОСНОВЕ НЕЙРОННОЙ СЕТИ

А.А. Макаренко, В.Т. Калайда

Томский государственный университет систем управления и радиоэлектроники
E-mail: makaleks@mail.ru

Предлагается метод и алгоритм локализации лица человека для автоматизированных систем распознавания и видеоконтроля на базе сверточных нейронных сетей. Преимущество использования свёрточных нейронных сетей в том, что они обеспечивают устойчивость к изменениям масштаба, смещениям, поворотам, смене ракурса и другим искажениям входного изображения. Обосновывается топология используемой нейронной сети и методика ее обучения.

Нейронные сети успешно применяют в решении многих проблем распознавания образов [1–3]: распознавание символов, распознавание объектов, и многих других. Проблема обнаружения образа лица очень трудна из-за большого разнообразия искажений, таких как различное выражение лица, условия съемки и т. д. Преимущество использования нейронных сетей для обнаружения лица – обучаемость системы для выделения ключевых характеристик лица из обучающих выборок.

В настоящее время наиболее часто в задачах распознавания и идентификации изображений используют классические нейросетевые архитектуры (многослойный персептрон, сети с радиально-базисной функцией и др.), но, как показывает анализ данных работ, применение классических нейросетевых архитектур к данной задаче является неэффективным по следующим причинам:

- к данной задаче обычно применяется ансамбль нейронных сетей (2–3 нейронные сети, обученные с различными начальными значениями синаптических коэффициентов и порядком предъявления образов), что отрицательно сказывается на вычислительной сложности решения задачи и соответственно на времени выполнения;
- как правило, классические нейросетевые архитектуры используются в совокупности с вспомогательными методами выделения сюжетной части изображения (сегментация по цвету кожи, выделение контуров и т. д.), которые требуют качественной и кропотливой предобработки обучающих и рабочих данных, что не является эффективным;
- нейросетевые архитектуры являются крайне чувствительными к влиянию различных внешних факторов (изменения условий съемки, присутствие индивидуальных особенностей на изображении, изменение ориентации).

Дополнительно возникают трудности применения традиционных нейронных сетей к реальным задачам распознавания и классификации изображений.

Во-первых, как правило, изображения имеют большую размерность, соответственно вырастает размер нейронной сети (количество нейронов и т. п.). Большое количество параметров увеличивает вместимость системы и соответственно требует большей обучающей выборки, что увеличивает время и вычислительную сложность процесса обучения.

Во-вторых, недостаток полносвязной архитектуры – то, что топология ввода полностью игнорируется. Входные переменные могут быть представлены в любом порядке, не затрагивая цель обучения. Напротив, изображения имеют строгую 2-мерную местную структуру: переменные (пиксели), которые являются пространственно соседними, чрезвычайно зависимы.

От данных недостатков свободны так называемые свёрточные нейронные сети. Свёрточные нейронные сети обеспечивают частичную устойчивость к изменениям масштаба, смещениям, поворотам, смене ракурса и другим искажениям. Свёрточные нейронные сети объединяют три архитектурных идеи, для обеспечения инвариантности к изменению масштаба, повороту, сдвигу и пространственным искажениям:

- локальные рецепторные поля (обеспечивают локальную двумерную связность нейронов);
- общие веса (обеспечивают детектирование некоторых черт в любом месте изображения и уменьшают общее число весовых коэффициентов);
- иерархическая организация с пространственными подвыборками.

Топология нейронной сети, используемой в работе, изображена на рис. 1.

Свёрточная нейронная сеть является многослойной. Используются слои двух типов: свёрточные и подвыборочные. Свёрточные и подвыборочные слои чередуются друг с другом. В свою очередь, каждый из этих слоёв состоит из набора плоскостей, причём нейроны одной плоскости имеют одинаковые веса (так называемые общие веса), ведущие ко всем локальным участкам предыдущего слоя (как в зрительной коре человека). Изображение предыдущего слоя сканируется небольшим окном и пропускается сквозь набор весов, а результат отображается на соответствующий нейрон текущего слоя. Таким образом, набор плоскостей представляет собой карты характеристик, и каждая плоскость находит «свои» участки изображения в любом месте предыдущего слоя.

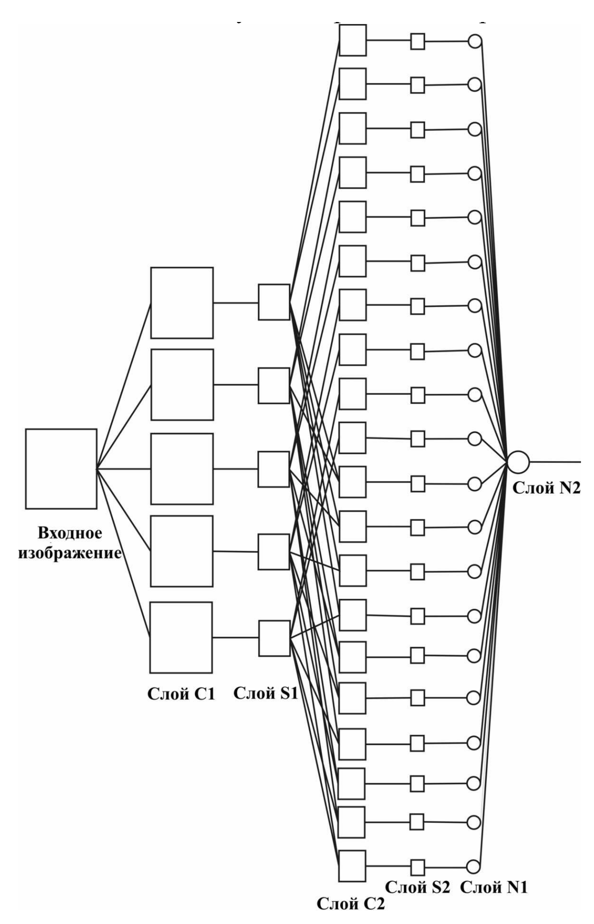


Рис. 1. *Схема используемой нейронной сети*

Используемая в работе нейронная сеть состоит из шести слоев. Входными данными нейронной сети являются полутоновые изображения размером 32×36 пикселей, которые классифицируются как лицо или «нелицо». Так как задача, решаемая нейронной сетью, – классификация, то для ее решения достаточно одного выхода. Выходное значение нейронной сети находится в интервале [–1;1], что соответственно означает отсутствие или присутствие лица на классифицируемом изображении.

Входной слой размером 32×36 нейронов не несет какой-либо функциональной нагрузки и служит лишь для подачи входного образа в нейронную сеть. Следом за входным слоем находится сверточный слой *C*1. Каждый нейрон в плоскости свёрточного слоя получает свои входы от некоторой области предыдущего слоя (локальное рецептивное поле), то есть входное изображение предыдущего слоя как бы сканируется небольшим окном и пропускается сквозь набор весов, а результат отображается на соответствующий нейрон свёрточного слоя.

Процесс функционирования нейрона свёрточного слоя задается выражением:

$$y_k^{(i,j)} = b_k + \sum_{s=1}^{K}\sum_{t=1}^{K} w_{k,s,t}\, x^{((i-1)+s,(j+t))},$$

где $y_k^{(i,j)}$ – нейрон *k*-ой плоскости свёрточного слоя, b_k – нейронное смещение *k*-ой плоскости, *K* – размер рецептивной области нейрона, $w_{k,s,t}$ – элемент матрицы синаптических коэффициентов, *x* – выходы нейронов предыдущего слоя.

Слой *C*1 состоит из 5 свёрточных плоскостей и выполняет свёртывание входного изображения с помощью синаптической маски размером 5×5, таким образом, слой *C*1 осуществляет 5 свёрток входного изображения.

Размер свёрточной плоскости определяется в соответствии со следующими выражениями:

$$w_c = w_u - K + 1,$$

$$h_c = h_u - K + 1,$$

где w_c, h_c – ширина и высота свёрточной плоскости соответственно, *w*, *h* – ширина и высота плоскости предыдущего слоя, *K* – ширина (высота) окна сканирования.

Исходя из выражений для расчета размера свёрточной плоскости имеем, что размер плоскости сверточного слоя *C*1 – 28×32 нейрона. Нейроны в слое организованы в плоскости, в пределах которых все нейроны имеют один и тот же набор синаптических коэффициентов. Набор выходных сигналов в такой плоскости называют картой характеристик. Полный свёрточный слой составлен из нескольких карт характеристик с различными наборами синапсов так, чтобы множественные характеристики могли быть извлечены в каждом местоположении. Таким образом, набор плоскостей представляет собой карты характеристик, и каждая плоскость находит «свои» участки изображения в любом месте предыдущего слоя.

Как указано выше каждая плоскость слоя *C*1 имеет собственную синаптическую маску и нейронное смещение, рецептивные области нейронов пересекаются, нейроны извлекают одни и те же особенности входного изображения, независимо от их точного местоположения. Таким образом, слой *C*1 имеет всего лишь 130 настраиваемых параметров (синапсов).

Следующий за слоем *C*1 подвыборочный слой *S*1 состоит из 5 карт характеристик и обеспечивает локальное усреднение и подвыборку. Этот слой также состоит из плоскостей количество плоскостей такое же, как и в предыдущем слое. Рецепторная область каждого нейрона – 2×2 область в соответствующей карте особенностей предыдущего слоя. Каждая нейрон вычисляет среднее его четырех входов, умножает на синаптический коэффициент, добавляет нейронное смещение и передает результат через активационную функцию. Процесс функционирования нейрона подвыборочного слоя задается следующим соотношением:

$$y_k^{(i,j)} = b_k + \frac{1}{4} w_k \sum_{s=1}^{2}\sum_{t=1}^{2} x^{(((i,j))+s,((i,j)))}.$$

Затем полученный результат подвыборки передается через активационную функцию. После операции подвыборки, точное местоположения и специфические признаки каждой особенности изо-

бражения становятся менее важными, что дает нейронной сети довольно большую степень инвариантности.

Смежные нейроны в подвыборочном слое имеют непересекающиеся рецептивные области. Следовательно, карта особенности слоя подвыборки имеет половину числа рядов и колонок карты особенности в предыдущем слое. В качестве активационной функции используется гиперболический тангенс $y = 1,7159\tanh\left(\frac{2}{3}x\right)$[4].

Каждая плоскость слоя *S*1 связана лишь с одной плоскостью слоя *C*1. Размер каждой плоскости слоя *S*1 – 14×16 нейронов, что вдвое меньше чем размер плоскости предыдущего слоя. Каждая плоскость слоя *S*1 имеет единственный синаптический коэффициент и нейронное смещение, что дает в итоге 10 настраиваемых параметров.

Свёрточный слой *C*2 состоит из 20 плоскостей, слои *S*1 и *C*2 перекрестно связаны. Плоскости слоя *C*2 формируются следующим образом: каждая из 5 плоскостей слоя *S*1 свёрнута 2 различными синаптическими масками 3×3, обеспечивая 10 плоскостей в *C*2, другие 10 плоскостей *C*2 получены, суммируя результаты 2 свёртываний на каждой возможной паре плоскостей слоя *S*1. Таким образом, сети добавляется способность объединять различные виды характеристик, чтобы составлять новые менее зависящие от искажений входного изображения.

Размер плоскости слоя *C*2 – 12×14 нейронов. Таким образом, данный слой имеет 290 синаптических коэффициентов. Слой *S*2 состоит из 20 плоскостей, размер каждой 6×7 нейронов. Каждая плоскость слоя *S*2 имеет единственный синаптический коэффициент и нейронное смещение, что дает в итоге 40 настраиваемых параметров.

Слои *N*1 и *N*2 содержат простые нейроны. Роль этих слоев состоит в обеспечении классификации, после того, как выполнены извлечение особенностей и сокращение размерности входа. В слое *N*1 находится 20 нейронов (по одному на каждую плоскость слоя *S*2), каждый нейрон полностью связан с каждым нейроном только одной плоскости слоя *S*2, он выполняет взвешенное суммирование своих 42 входов, добавляет нейронное смещение и пропускает результат через активационную функцию. Таким образом, данный слой содержит 860 синаптических коэффициентов.

Единственный нейрон слоя *N*2 полностью связан со всеми нейронами слоя *N*1. Роль этого нейрона в вычислении окончательного результата классификации. Выход этого нейрона используется для классификации входного образа на лица и не лица.

Использование принципа объединения весов дает эффект уменьшения количества настраиваемых параметров нейронной сети. Данная нейронная сеть имеет 1351 синаптический коэффициент.

Способность к обучению является фундаментальным свойством мозга. В контексте искусственных нейронных сетей процесс обучения может рассматриваться как настройка архитектуры сети и весов связей для эффективного выполнения специальной задачи. Процесс функционирования нейронной сети зависит от величин синаптических связей, поэтому, задавшись определенной структурой нейронной сети, отвечающей какой-либо задаче, необходимо найти оптимальные значения всех переменных коэффициентов (некоторые синаптические связи могут быть постоянными). Этот этап называется обучением нейронной сети, и от того, насколько качественно он будет выполнен, зависит способность сети решать поставленные перед ней проблемы во время эксплуатации. В основе всех алгоритмов обучения положен единый принцип – минимизация эмпирической ошибки. Функция ошибки, оценивающая данную конфигурацию сети, задается извне в зависимости от того, какую цель преследует обучение. Но далее сеть начинает постепенно модифицировать свою конфигурацию – состояние всех своих синаптических весов таким образом, чтобы минимизировать эту ошибку.

Для обучения описанной нейронной сети был использован алгоритм обратного распространения ошибки (*back propagation*). Метод был предложен в 1986 г. Румельхартом, Макклеландом и Вильямсом [5].

Обучение сети начинается с предъявления образа и вычисления соответствующей реакции. Сравнение с желаемой реакцией дает возможность изменять веса связей таким образом, чтобы сеть на следующем шаге могла выдавать более точный результат. Обучающее правило обеспечивает настройку весов связей. Информация о выходах сети является исходной для нейронов предыдущих слоев. Эти нейроны могут настраивать веса своих связей для уменьшения погрешности на следующем шаге.

Когда ненастроенной сети предъявляется входной образ, она выдает некоторый случайный выход. Функция ошибки представляет собой разность между текущим выходом сети и идеальным выходом, который необходимо получить. Для успешного обучения сети требуется приблизить выход сети к желаемому выходу, т. е. последовательно уменьшать величину функции ошибки. Это достигается настройкой межнейронных связей. Каждый нейрон в сети имеет свои веса, которые настраиваются, чтобы уменьшить величину функции ошибки. Выражение для коррекции синаптических коэффициентов имеет вид:

$$w_{ij}(t+1) = w_{ij}(t) + \eta \frac{dE}{dw_{ij}}, \quad (1)$$

где $w_{ij}(t)$ и $w_{ij}(t+1)$ – вес связи между *i*-м и *j*-м нейронами на текущем и последующем шаге обучения, $\frac{dE}{dw_{ij}}$ – производная функции ошибки, η – параметр скорости обучения.

В основе алгоритма обратного распространения ошибки лежит методика, позволяющая быстро вычислять вектор частных производных (градиент) сложной функции многих переменных, если структура этой функции известна. В качестве такой функции в алгоритме рассматривается функция ошибки сети и учитывается тот факт, что структура функции ошибки сети полностью определяется архитектурой нейронной сети, которая считается известной.

Метод обратного распространения ошибки может быть очень медленным особенно для многослойных сетей, где поверхность целевой функции является неквадратичной, невыпуклой и высоко размерной с множеством локальных минимумов и/или плоских областей.

Чтобы вычислить градиент на каждой итерации, ур. (1) должно быть рассчитано для всего обучающего набора данных. Эта процедура называется пакетным обучением; при таком способе обучения вся обучающая выборка рассматривается до обновления синаптических коэффициентов. Альтернативно, можно использовать стохастическое обучение (онлайн), где отдельный пример $\{Z_l,D_l\}$ выбран (например, случайно) из учебного набора на каждой итерации. Поскольку эта оценка градиента является зашумленной, синаптические коэффициенты, возможно, не перемещаются точно по градиенту на каждой итерации. Как далее будет показано, этот «шум» на каждой итерации может быть полезным. Стохастическое обучение – более предпочтительный метод для алгоритма обратного распространения ошибки: результат достигается обычно намного быстрее, чем при пакетном методе.

Стохастическое обучение часто приводит к лучшим решениям из-за шума в обновлениях весовых коэффициентов. Нелинейные сети обычно имеют множество локальных минимумов различной глубины. Цель обучения состоит в том, чтобы определить местонахождение одного из этих минимумов. Пакетное обучение обнаружит минимум любого бассейна, куда первоначально помещены синаптические коэффициенты. В стохастическом обучении шумовая составляющая в обновлениях может привести к весам, вскакивающим в бассейн другого, возможно более глубокого локального минимума. Результаты экспериментов [6] говорят о том, что затраты на одномерную оптимизацию шага не дают практической пользы, обучение по суммарному градиенту (с оптимизацией шага или без неё) всегда проигрывает стохастическому обучению.

Нелинейные активационные функции – то, что придает нейронным сетям их нелинейные свойства. В данной работе в качестве активационной функции используется гиперболический тангенс. Это обусловлено следующими причинами:

- симметричные активационные функции, типа гиперболического тангенса, обеспечивают более быструю сходимость, чем стандартная логистическая функция;
- данная функция имеет простую и непрерывную первую производную.

В задачах классификации целевые выходы типично являются бинарными (например, ±1), т. е. целевые выходы устанавливаются по асимптотам активационной функции. Такой подход имеет несколько недостатков.

1. Результат обучения может быть нестабилен. Процесс обучения будет настраивать выход нейронной сети как можно близко к целевым значениям, которые могут быть достигнуты только асимптотически. В результате синаптические коэффициенты (выходного слоя и скрытых слоев) стремятся к большим и большим значениям, в которых производная активационной функции близка к нулю. В результате коррекция синаптических коэффициентов может стать незначительной.
2. В случае если выходы насыщенные, сеть не дает уверенности в правильности классификации. Когда входной образ находится на границе решения, результат классификации сомнителен. Бо́льшие значения весовых коэффициентов имеют тенденцию смещать выходное значение к хвостам активационной функции независимо от класса. Таким образом, сеть может предсказать неправильный класс, не давая уверенности в результате. Решение этих проблем состоит в том, чтобы заставить целевые выходы быть в пределах диапазона активационной функции. Установка целевых значений на грани максимума второй производной активационной функции – лучший способ использовать в своих интересах нелинейность, не насыщая активационную функцию. По этой причине используется указанная функция активации. Она имеет максимальную вторую производную в точках ±1, которые соответствуют целевым значениям, используемым в задачах классификации.

Начальная инициализация синапсов нейронной сети имеет огромное влияние на количество итераций обучения. От того, насколько удачно выбраны начальные значения синаптических коэффициентов зависит, как долго сеть за счет обучения и подстройки будет искать их оптимальные величины, и найдет ли она их. Начальные значения синаптических коэффициентов могут существенно влиять на процесс обучения. Синаптические коэффициенты должны быть выбраны случайно, но таким способом, чтобы активационная функция прежде всего активизировалась в своей линейной области. Как показано в работе [4] для достижения данного эффекта необходимо согласование между нормализацией входных значений нейронной сети, выбором активационной функции и выбором начальных значений весовых коэффициентов. Таким образом, для выбранной активационной функции и выбранного метода нормализации входных данных в данной работе синаптические коэффициенты инициализировались следующим

образом: значения весовых коэффициентов были выбраны случайным образом из нормального распределения с нулевым средним и стандартным отклонением

$$\sigma_w = \sqrt{m},$$

где m – число связей, входящих в нейрон.

В выражении (1) параметр η по сути является мерой точности обучения сети. Чем он больше, тем более грубым будет следующее уменьшение суммарной ошибки сети. Чем он меньше, тем больше времени сеть будет тратить на обучение и тем более возможно ее попадание в окрестность локального минимума. Интуитивно понятно, что выбор индивидуальной скорости обучения для каждого синаптического коэффициента может ускорить сходимость и улучшить качество решения (в зависимости от формы поверхности функции ошибки, некоторые веса могут требовать небольшой скорости обучения, чтобы избежать расхождения, в то время как другие могут требовать большого значения скорости обучения, чтобы ускорить схождение алгоритма). Поэтому в данной работе каждому синаптическому коэффициенту дана индивидуальная скорость обучения. Для автоматической подстройки скорости обучения на каждой итерации обучения использовался алгоритм, предложенный в работе [7].

Как входами, так и выходами могут быть совершенно разнородные величины. Очевидно, что результаты нейросетевого моделирования не должны зависеть от единиц измерения этих величин. А именно, чтобы сеть трактовала их значения единообразно, все входные и выходные величин должны быть приведены к единому масштабу. Скорость сходимости обычно быстрее, если среднее значение каждой входной переменной по учебному набору близко к нолю. Поэтому, нужно отмасштабировать входные значения так, чтобы среднее число по учебному набору было близко к нолю [4]. Таким образом, в данной работе все входные значения из интервала [0; 255] (т. е. цвет пиксела от черного до белого) были предварительно отмасштабированы в интервал [–1; 1].

В качестве учебного набора используется большая коллекция изображений, полученных из различных источников. Эта коллекция эффективно охватывает изменчивость и богатство естественных данных, чтобы обучить данную систему для работы в реальных условиях. В данной работе, используется окно размером 32×36 пикселей, содержащее лицо и некоторую часть фона, таким образом, добавляя к входному окну некоторую вспомогательную информацию: границу лица и некоторую часть фона. В процессе извлечения лиц не выполнялось какой-либо нормализации изображений типа выравнивания гистограммы или коррекции яркости. Кроме того, примеры лиц не нормализовывались так, чтобы глаза, рот и другие части лиц всегда оставались примерно на том же самом положении. Кроме того, как упомянуто ранее, сверточная нейронная сеть является устойчивой к изменению масштаба и положения, таким образом, для повышения этой надежности необходимо давать нейронной сети ненормализованные примеры. Кроме того, для создания большого количества примеров и увеличения инвариантности к небольшим вращениям и изменениям в интенсивности, к выше упомянутому набору применялся ряд преобразований, включая отражение, вращение до ±20°. Некоторые из изображений обучающей выборки представлены на рис. 2.

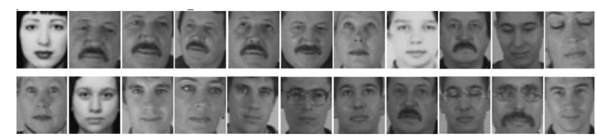


Рис. 2. *Изображения из лицевой обучающей выборки*

Сбор представительного набора нелицевых образов более труден, т. к. фактически, любое изображение может принадлежать к классу нелицевых образов. Практическое решение этой проблемы состоит в стратегии самонастройки [8], при которой система многократно переобучается на ряде изображений пейзажа, которые не содержат лиц. Перед началом процедуры самонастройки, был построен начальный учебный набор. Большинство из этих изображений содержит части лиц, поскольку как было замечено в ранних экспериментах, этот вид изображений является серьезным источником ложных сигналов. Некоторые из этих изображений показаны на рис. 3.



Рис. 3. *Изображения из стартовой нелицевой обучающей выборки*

Процесс обнаружения лица состоит из 3 этапов (рис. 4):

1. Входное изображение последовательно масштабируется, получается пирамида изображений, затем каждое изображение сканируется нейронной сетью, происходит выделение участков-кандидатов.
2. Все участки-кандидаты приводятся к масштабу входного изображения, и затем смежные участки-кандидаты группируются в кластеры, устраняя избыточность.
3. Далее происходит дальнейшая проверка кластеров, выносится решение о принадлежности или не принадлежности каждого участка к классу лиц.

Нами было произведено тестирование данной программной системы на наборе из 120 изображений. На данном тестовом наборе получена точность обнаружения равная 97 %. Исходя из этого, считаем, что предложенная нами топология сверточной нейронной сети обеспечивает необходи-

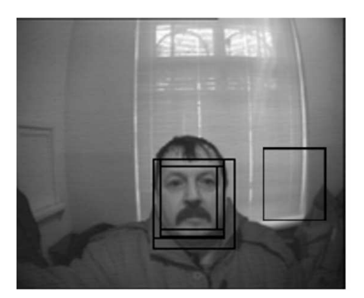




Рис. 4. *Процесс обнаружения лица*

мую надежность и инвариантность к искажениям и зашумлениям входного сигнала, что позволяет создать на ее базе систему выделения сюжетной части изображения. Модификации процесса обучения обеспечивают качественный процесс обучения нейронной сети, улучшают ее обобщающие и классифицирующие способности, позволяют использовать данную программную систему для решения практических задач в системах видеонаблюдения и контроля доступа.

Работа выполнена при поддержке РФФИ, проект № 06-08-00751.

СПИСОК ЛИТЕРАТУРЫ

1. Moody J., Darken C.J. Fast Learning In Netwoks of Locally Tuned Processing Units // Neural Computation. – 1989. – № 1. – P. 281–284.
2. Haddadnia J., Faez K. Human face Recognition Using Radial Basis Function Neural Network // Proc. of 3rd International Conf. on Human and Computer. – Aizu, Japan, 6–9 Sept. 2000. – P. 137–142.
3. Lin S.H., Kung S.Y., Lin L.J. Face Recognition/Detection by Probabilistic Decision-Based Neural Network // IEEE Trans. Neural Networks. – 1997. – V. 8. – № 1. – P. 114–132.
4. LeCun Y. Generalization and network design strategies // Proc. of the Intern. Conf. Connectionism in Perspective. – University of Zurich, 10–13 October 1988. – P. 143–155.
5. Rumelhart D.E., Hinton G.E., Williams R.J. Learning internal representations by error propagation // Parallel distributed processing. – 1986. – V. 1. – P. 318–362.
6. Царегородцев В.Г. Общая неэффективность использования суммарного градиента выборки при обучении нейронной сети // Нейроинформатика и ее приложения: Матер. XII Всеросс. семинара. – Красноярск, 2004. – С. 145–151.
7. Murata N., Miller K.R., Ziehe A., Amari S. Adaptive on-line learning in changing environments // Advances in Neural Information Processing Systems. – 1997. – V. 9. – P. 599–623.
8. Sung K.K. Learning and Example Selection for Object and Pattern Detection: PhD thesis. – MIT AI Lab, 1996. – 150 p. (ftp://publications.ai.mit.edu/ai-publications/1500-1999/AITR-1572.ps.Z)

УДК 004.93'12

МЕТОДИКА ПОСТРОЕНИЯ СИСТЕМ РАСПОЗНАВАНИЯ АВТОМОБИЛЬНОГО НОМЕРА

А.И. Елизаров, А.В. Афонасенко

Томский государственный университет систем управления и радиоэлектроники
E-mail: alex@iao.ru

Рассмотрена технология создания системы поиска и распознавания автомобильных номеров. Предлагаются модифицированные алгоритмы, позволяющие повысить надежность и точность подобных систем. Данные алгоритмы использовались для создания программного комплекса, состоящего из набора программных модулей, выполняющих предварительную обработку изображений, обнаружение номерной пластины, извлечение символов, распознавание символов.

Проблема автоматизированного оперативного распознавания текстовой информации является актуальной задачей, связанной с широким классом практических приложений. Одной из таких задач является распознавания автомобильных номеров. Создание автоматической системы, регистрирующей автомобильные номера, позволяет:

- автоматизировать контроль въезда и перемещения транспортных средств на объектах с ограниченным доступом и закрытых территориях;
- отслеживать въезд и выезд на автостоянках, осуществлять автоматический подсчет стоимости предоставленных услуг, контролировать свободное место;
- автоматизировать контроль выезда оплаченных или неоплаченных транспортных средств на станциях технического обслуживания и автокомбинатах, контролировать загрузку зоны обслуживания;



Рис. 1. *Общая структура типовой системы распознавания автомобильных номеров*

- отслеживать въезд, выезд и время нахождения транспортных средств на территории склада и терминала, предотвращать возможные хищения;
- на автомагистралях обеспечить контроль транспортных потоков и осуществлять автоматическое трассирование угнанных транспортных средств и тех, за которыми числятся правонарушения;
- автоматизировать сбор статистики для муниципальных служб.

На сегодняшний день в мире уже существует несколько систем автоматического распознавания автомобильных номеров. Все они далеки от совершенства и постоянно модифицируются. Тем не менее, общая структура и решения данной задачи уже сформированы (рис. 1).

Устройство съема изображения – видеокамера. На установку видеокамеры накладывается ряд ограничений. Оптическое разрешение камеры выбирается таким образом, чтобы номерная пластина занимала в кадре по горизонтали от 25 до 33 %. Максимально допустимая скорость автомобиля, при которой возможно распознавание, зависит от угла наклона камеры к полотну дороги. В работе [1] декларируется: «Типовой угол наклона камеры должен равняться 40°, чтобы впереди идущий автомобиль не загораживал следующий», что также накладывает ряд ограничений. Вертикальный размер номерной пластины уменьшается пропорционально значению косинуса угла наклона камеры. Следовательно, необходимо увеличивать оптическое разрешение, что приводит к сокращению ширины поля зрения. Некоторые номерные пластины могут быть установлены с наклоном в сторону дорожного полотна. При большом угле наклона распознавание таких номеров весьма проблематично. Оптимальной считается горизонтальная установка видеокамеры, т. е. на том уровне, где находиться номерная пластина. Такая установка используется в системах автоматического пропуска на закрытую территорию, парковки и т. д.

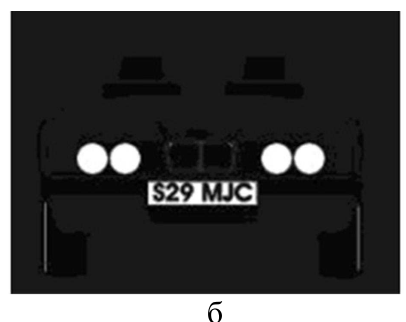
Другим важным параметром является минимально допустимая контрастность изображения номерной пластины. В некоторых системах устанавливается дополнительный модуль, состоящий из инфракрасной подсветки и соответствующего фильтра [2]. Такой подход позволяет повысить контрастность номерной пластины по отношению к остальному изображению. Эта технология основана на том, что номерная пластина имеет специальное светоотражающее покрытие, при котором отраженный свет распространяется в обратном направлении распространению падающего света (т. е. угол между падающим и отраженным лучом равняется 0°), вследствие чего камерой будет восприниматься в основном инфракрасный свет и свет, отраженный от номерной пластины (рис. 2). Снимок в этом случае, будет монохромным без деталей, за исключением номерной пластины.

Предобработка полученного изображения включает следующие этапы.

1. Коррекция изображения – эквилизация, ограничение экстремальных значений яркости, видоизменение гистограммы распределения яркостей.
2. Устранение эффекта смазывания изображения, возникающего в связи с тем, что скорость автомобиля больше чем скорость регистрации (компенсация сдвига).
3. Устранение избыточной информации – использование инфракрасной (ИК) подсветки, бинаризация, разбиение изображения на отдельные цветовые регионы.
4. Использования программного детектора движения для локализации автомобиля на изображении.



а



б



в

Рис. 2. *Предобработка изображения: а) оригинальное изображение, б) с применением ИК-подсветки, в) бинаризованное изображение*

Недостатком предобработки с бинаризацией является то, что выбираемый порог бинаризации не обеспечивает необходимого качества для любого типа изображения. Такие факторы, как освещение или даже цвет транспортного средства влияют на качество бинаризации изображения. Использование методов адаптивной бинаризации позволяет, тем не менее, решить данную проблему более качественно.

Одним из применяемых подходов для локализации номерной пластины является метод обнаружения границ изображения с использованием таких алгоритмов как алгоритм Собеля, Кенни, Робинсона и др. [3]. Алгоритм обнаружения границ должен обнаруживать как горизонтальные, так и вертикальные края. Результирующее изображение, после выделения краев должно содержать большое количество линий в области пластины номерного знака, т. к. она содержит символы. Это основное свойство, которое используется для выделения области пластины номерного знака на изображении.

Для локализации области номерного знака создается окно, приблизительно равное размеру пластины номерного знака на изображении. Это окно используется, чтобы оценить количество граней во всех областях изображения, имеющих наибольший контраст. Окно накладывается на результирующее изображение в наиболее контрастных зонах. Если количество граней находится в заданном диапазоне, то данная область отмечается как область, возможно содержащая номерной знак. Необходимое количество граней определяется экспериментально. Результат работы этого окна – список возможных областей – кандидатов, в которых может содержаться пластина номерного знака. Основным недостатком этого подхода является то, что процесс выбора кандидата является медленным, поскольку значения всех пикселей в выделенном окне должны быть неоднократно суммированы. Другим недостатком является набор большого количества кандидатов в номерные знаки.

Альтернативным подходом для обнаружения номерной пластины является преобразование Хафа, которое используется для обнаружения участков различных форм в изображении, таких как круг, эллипс, прямая и т. д. [4]

На вход преобразования Хафа подается бинарное изображение с вертикальными и горизонтальными выделенными краями. На выходе преобразования получаем набор прямых линий, ограничивающих предполагаемый номерной знак, то есть список возможных кандидатов на местоположение пластины номерного знака.

Алгоритм преобразования Хафа для обнаружения прямой линии выглядит следующим образом.

1. Выбор начального пикселя $A(x,y)$.
2. Выбор конечного пикселя $B(x,y)$.
3. Подсчет точек бинарного изображения по линии AB.
4. Если количество подсчитанных пикселей больше, чем заданное пороговое значение, то линия AB присутствует в изображении и помечается.
5. Возврат к шагу 1 и выбор двух других пикселей, пока не достигнута последняя точка изображения.

Горизонтальные и вертикальные пары линий сопоставляются. Горизонтальные и вертикальные пары линий, составляющие прямоугольник с отношением сторон примерно равных отношению сторон номерной пластины, отмечаются как область, возможно содержащую номер.

Одним из недостатков преобразования Хафа является тот факт, что вертикальные линии на номерном знаке значительно короче горизонтальных и, следовательно, могут быть более зашумленными.

После того как номерная пластина локализована, выполняется операция обнаружения символов. Для устранения избыточной информации используется алгоритм бинаризации. Выбор порога в алгоритме осуществляется на основе яркостной гистограммы изображения, которая представляет собой одномерный массив $H[0,255]$, в каждой ячейке которого содержится число точек изображения, имеющих значение интенсивности I.

Согласно наблюдениям, автомобильный номер обладает следующим свойством: средняя площадь всех символов составляет около 23 % от площади всего номера, который имеет форму прямоугольника. Для разных номеров отклонения от этого значения не превышают 5 %. В этом случае порог бинаризации T может быть определен следующим выражением $\sum_{j=0}^{T-1} H[j] \le 0,23S$, где S – площадь прямоугольника, описывающего номер автомобиля. Результат работы данного алгоритма представлен на рис. 3.

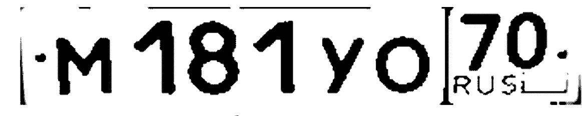


Рис. 3. *Бинарное изображение*

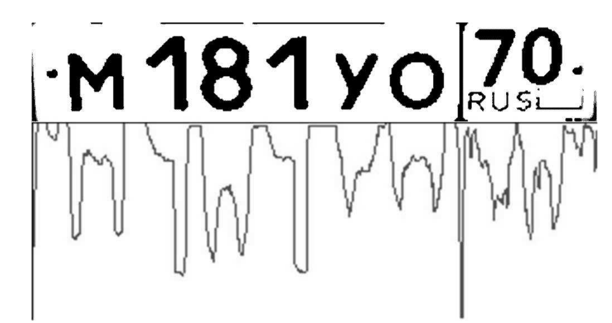


Рис. 4. *Горизонтальная проекция (распределение средней интенсивности) номерной пластины*

Следующим шагом является поиск отдельных символов. Для этой цели мы используем метод, основанный на построении проекции средней интенсивности (рис. 4). Суть этого подхода сводится к следующему,

вычисляется средняя интенсивность в каждом столбце, и в тех местах, где нет символа, средняя интенсивность будет значительно отличаться. Далее выполняя ту же операцию по строкам, получается набор отдельных символов, которые уже можно распознавать.

После выполнения данной операции мы можем получить как символы, так и различные помехи, которые необходимо устранить. Это достигается путем проверки ряда условий, которым должен удовлетворять символ как геометрический объект.

Во-первых, вычисляются размеры прямоугольника, описывающего область-кандидат. Проверяется выполнение условия $0,9 \le \frac{b}{a} \le 3,5,$ где a и b – соответственно ширина и высота объекта. Объекты, не удовлетворяющие данному соотношению, отбрасываются из рассмотрения.

Во-вторых, проверяется выполнение условия $0,4 \le \frac{b}{h} \le 0,9,$ где h – высота прямоугольника, описывающего номер автомобиля. Это соотношение накладывает ограничение на минимальную и максимальную высоту символов по отношению к высоте номера.

Наконец, третье условие имеет целью удаление мелких областей, не представляющих интереса. Среди объектов, удовлетворяющих вышеописанным условиям, выбирается объект с максимальной площадью описывающего прямоугольника $S_{max}=a \cdot b$. Затем среди оставшихся областей выбираются те, которые удовлетворяют условию $\frac{S}{S_{max}} \ge 0,1,$ где S – площадь прямоугольника объекта-кандидата.

Проверка этих несложных условий позволяет устранить все лишние объекты, выделив только восемь знаков номера (рис. 5).

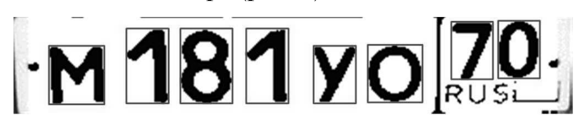


Рис. 5. *Выделенные символы*

Поскольку автомобильный номер содержит ограниченный набор символов фиксированного шрифта, в данном случае целесообразно применение шрифтового алгоритма распознавания. Принцип действия основан на прямом сравнении изображения символа с эталоном. Степень несходства при этом вычисляется как количество несовпадающих пикселей. Для обеспечения приемлемой точности шаблонного метода требуется предварительная обработка изображения: нормализация размера, наклона и толщины штриха. Эталон для каждого класса обычно получают, усредняя изображения символов обучающей выборки.

Этот метод прост в реализации, работает быстро, устойчив к случайным дефектам изображения, однако имеет относительно невысокую точность. Широко используется в современных системах распознавания символов.

Для распознавания символов нами используется одна из модификаций шрифтового алгоритма – метод зон [6] . Данный метод основан на сравнении зонного представления изображения символа с эталонным набором зонных представлений символов. Поэтому алгоритм распознавания состоит из трех основных процедур: обучение, загрузка набора эталонных зонных описаний, распознавание.

На этапе обучения использовался полный эталонный набор бинарных изображений символов, изображения всех допустимых символов. При этом для каждого эталонного изображения в наборе выполнялись следующие действия.

- Определение минимального прямоугольника, содержащего все черные пиксели.
- Равномерное разбиение рамки на $N \times M$ прямоугольных зон.
- Подсчет числа пикселей, принадлежащих каждой зоне.
- Формирование вектора зонного описания, состоящего из значений числа черных пикселей для каждой зоны, нормированных путем деления на суммарное число черных элементов всего изображения.

Подсчет числа пикселей, принадлежащих каждой зоне, осуществляется путем голосования черных пикселей. При этом каждый пиксель рассматривается как квадрат размером 1×1, положение границ зон рассчитывается с субпиксельной точностью, и каждый черный пиксель голосует в пользу тех зон, с которыми он пересекается, с весом, равным площади пересечения.

В результате этапа обучения для каждого символа формируется файл зонных описаний, имя символа в ASCII-коде и вектор зонного описания размера $N \times M$ с элементами типа *float*. Вектор зонного описания записывается по строкам, слева направо сверху вниз.

При загрузке эталонных данных происходит считывание эталонного файла и формирование соответствующего списка эталонных векторов зонных описаний.

При анализе каждого символа производятся следующие операции:

1. формирование вектора зонного описания;
2. формирование вектора расстояний;
3. классификация символа на основе вектора расстояний.

Формирование вектора зонного описания осуществляется, как было описано ранее для эталонных изображений.

Формирование вектора расстояний предполагает последовательное вычисление евклидовых расстояний между вектором зонного описания тестируемого символа и векторами зонных описаний каждого эталона в списке. Классификация символа осуществляется путем анализа вектора расстояний и выбора «ближайшего соседа» среди эталонных зонных описаний.

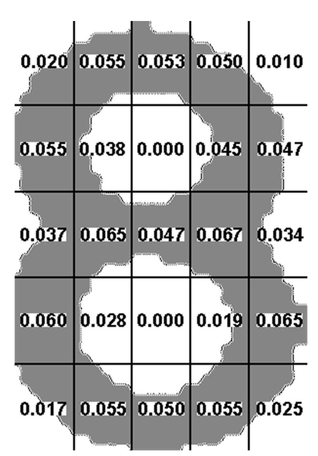


Рис. 6. *Эталонное изображение символа «8»*

С учетом структуры штрихов начертания символов, для размера изображений символов порядка 10×15 пикселей оптимальным числом зон будет $N \times M$, где N=5 и M=5. При таком выборе зонного разрешения использованный способ нормирования вектора зонного описания позволяет обеспечить устойчивость данного метода распознавания к изменению толщины штрихов символа вследствие нестабильности яркостных характеристик изображения.

На рис. 6 показано эталонное изображение символа «8», разбиение на зоны и пиксельные веса каждой зоны.

Описанная выше, последовательность алгоритмов была использована при создании программы поиска и распознавания автомобильных номеров. Программа разработана в среде Delphi 7 и протестирована на 200 изображениях, вероятность распознавания автомобильного номера составила 85 %.

Работа выполнена при поддержке РФФИ, проект № 06-8-00751.

СПИСОК ЛИТЕРАТУРЫ

1. Department of Electrical and Computer Web Pages. License Plate Recognition Final Report. Обращение к документу 19.08.2006. http://www.ece.cmu.edu/~ee551/Final_Reports/Gr18.551.S00.pdf
2. Kim K.I., Kim K.K., Park S.H., Jung K., Park M.H., Kim H.J. VEGA VISION: A Vision System For Recognizing License Plates // IEEE Intern. Symp. on Consumer Electronics (ISCE 99). – 1999. – V. 2. – P. 176–181.
3. Прэтт У. Цифровая обработка изображений. Пер. с англ. – М.: Мир, 1982. – Т. 1, 2.
4. Gonzalez R.C., Woods R.E. Hough Transform Edge Linking // Digital Image Processing. – 2001. – № xx. – P. 569–577.
5. Уоссермен Ф. Нейрокомпьютерная техника. – М.: Мир, 1992. – 184 с.
6. M.V. Keldysh Institute for Applied Mathematics Russian Academy of Sciences Web Pages. Алгоритмическое обеспечение прототипа устройства считывания паспортов и виз. Обращение к документу 05.07.2006. http://www.keldysh.ru/papers/2003/prep46/prep2003_46.html

УДК 004.932.4